UNIVERSITY OF CALIFORNIA, DAVIS

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO



SANTA BARBARA • SANTA CRUZ

DAVID M. ROCKE, PROFESSOR
DIVISION OF BIOSTATISTICS AND DEPARTMENT OF APPLIED SCIENCE
UNIVERSITY OF CALIFORNIA, DAVIS
ONE SHIELDS AVE
MED SCI 1-C, ROOM 145
DAVIS, CA 95616-8638

PHONE: (530) 752-6999 FAX: (530) 754-6793 EMAIL: DMROCKE@UCDAVIS.EDU HTTP://WWW.IDAV.UCDAVIS.EDU/~DMROCKE

April 26, 2007

Steven J. Brisby, Manager Fuels Section, Stationary Source Division California Air Resources Board 1001 "I" Street P.O. Box 2815 Sacramento, CA 95812

Dear Steve:

I am writing with comments on a variety of issues that you have asked me to examine regarding the revisions of the predictive model.

I have checked the ARB staff's basic SAS programming and have found no issues of error or inconsistent logic.

The staff decided not to use RVP interactions in the model due to the high correlation of RVP with other fuel properties such as T50 and T90 and since RVP was not sufficiently varied internal to most of the studies. I have examined this decision and believe that it is reasonable and supported by the available evidence.

In the Tech 4 NOx model, two quadratic terms were removed (olefins and aromatics) because engineering judgment and comparison with previous work suggested that these effects were spurious and due to incidental correlations. I have examined the data and results and conclude that these changes were statistically justifiable and represent good practice.

There is a proposal to model high-emitters separately in one case. NOx emissions would be modeled with one model for automobiles whose average NOx emissions were higher than a certain threshold and another model for the case where NOx emissions were below that threshold. I find this proposal problematic on a number of fronts and recommend against its adoption.

The CARB Predictive Model has always been developed as a compromise between detail and simplicity. No one believes that the true response of automobiles to fuel changes can be exactly described by a quadratic in the specified fuel parameters. Instead, the idea is that the model should be faithful to the data and as easy to use as is consistent with that end. Given the large number of automobiles in the data set, there seems no doubt that statistical tests could be produced showing that a variety of complications of the model would fit "better" than the base model. This might include cubic

terms or estimated data transformation parameters after the fashion of Box and Tidwell. At some point the complexity needs to be reined in to make the model useful. In my opinion, the division of automobiles into emissions classes with no engineering basis is not warranted, even if a statistical test suggests that the model fit is better. This is particularly true in this case, since the "best" division is not into normal and high emitters, but rather into below-average and above-average emitters, which seems to make no engineering sense. I have examined the comments by Gary Herwick on a possible engineering basis for such a bifurcation. I am not an expert on automotive technology, but I am experienced in the issues relating natural variability and a regulatory limit to potential settings of average values. I find it statistically plausible that average emissions would be set to some fraction of the regulatory limit both to accommodate natural variation and to allow for drift and aging. This would imply, though, that the range of emissions from properly functioning automobiles would be substantial and that one would not expect any discontinuity at the average emissions level as the cars left the factory. Further, if there was such a discontinuity, one would expect to see it in several pollutants at this below-limits level, whereas the cutoff estimated by maximum likelihood varies substantially between the models.

A second problem with the bifurcation method, is that it introduces a needless discontinuity in the model. At present, predictions are based on continuous fuel parameters, in which the model predictions vary continuously with the parameters. There are separate models for emissions technology classes, which is appropriate considering the engineering differences, but there is no ambiguity: the year of an automobile's manufacture is not subject to dispute. When the model is to vary depending on a sharp division between the average emissions levels of different vehicles, and when those models may be very different, we are making the assertion that there is a sudden jump in emissions behavior at a certain point. I know of no support for this in the studies we are using or elsewhere in the literature.

In the bifurcation method that was proposed, the division between the two models is based on the observed response for a given automobile to NOx tests on one or more fuels. The NOx response forms the left-hand-side, or dependent variable in the regression, and the bifurcation proposes to use a function of these response values to form a new variable that will then be used as an independent variable on the right-hand-side of the regression in the form of interaction terms. In general, the use of 'independent' variables that are derived from the dependent variable is problematic. This defect may be somewhat ameliorated in this case due to the fact that the indicator variable is based on potentially several responses, and the fact that the car-to-car differences are relatively large compared to the other factors, but it is still a procedure that does not have good theoretical grounding. Further work by Jonathan Cohen has ameliorated this particular problem by examining a principled resampling-based method of setting p-values. However, the other problems still remain too large for me to concur with this idea.

Comments on decisions of the Statistics Working Group:

Averaging the repeats but not the true replicates makes sense, due to the lower variability of contiguous repeats.

Removing points as outliers only with specific scientific or engineering justification is sound practice, in my opinion.

Constructing the Tech 5 model as offsets to the Tech 4 model makes sense due to the relatively small database for Tech 5, and due to the fact that differences between Tech 4 and Tech 5 are likely smaller than differences between Tech 4 and previous technologies.

Sincerely,

David M. Rocke

David M Roche